

# Principes de conception d'un dépôt central de métadonnées comme élément clé d'un système intégré d'information sur la santé

## Messages clés

### *Le défi*

Le système de santé suisse (SSM) est un système distribué et complexe avec divers éléments en interaction, impliquant cinq groupes majeurs d'acteurs (population, prestataires de santé, payeurs, ressources humaines et gouvernance) pour lesquels les données sont collectées et analysées à des fins distinctes. Un système national intégré d'information sur la santé (SIS) est conçu avec l'objectif de générer des informations pour améliorer les décisions de gestion de la santé à tous les niveaux, sur la base de données fiables et opportunes. Afin d'atteindre l'objectif global d'intégration des données, un système national suisse d'information sur la santé (*Swiss health information system*, SHIS) est confronté au défi d'identifier et d'accéder à des données provenant de diverses sources hétérogènes afin de fournir des informations pertinentes. Il est donc nécessaire d'identifier une solution de gestion des données pour aider à gérer l'hétérogénéité et la dispersion des données de santé suisses.

Dans le contexte d'un SHIS, deux approches différentes peuvent être envisagées pour une infrastructure d'intégration des données appropriée : une approche centralisée (entrepôt de données et/ou lac de données) et une approche décentralisée (bases de données fédérées utilisant des services Web). L'identification d'une approche optimale pour la Suisse doit tenir compte des aspects techniques, ainsi que des aspects juridiques (par exemple, la protection des données), culturels (par exemple, la vie privée) et politiques (par exemple, la fédéralisation) pertinents pour le contexte suisse.

### *Options pour relever le défi*

Compte tenu des contraintes actuelles, cette lettre propose une infrastructure hybride qui prend en charge les modes d'accès aux données suivants : (1) mode d'accès par pont : pour les données stockées dans un entrepôt de données central, (2) mode d'accès par ferryboat : pour les données stockées dans des référentiels locaux persistants, et (3) mode d'accès par grue : pour les données stockées dans des bases de données locales temporelles.

Le défi soulevé par une infrastructure de données hybride, sans entrepôt de données central, est la capacité de récupérer les informations sur les ensembles de données disponibles. Par conséquent, une infrastructure de gestion des métadonnées est cruciale pour garantir que les données nécessaires sont récupérées et accessibles à partir des bonnes bases de données, en respectant les spécificités de chaque ensemble de données. L'élément principal d'une telle infrastructure est un dépôt central de métadonnées désigné comme le principal point d'accès pour l'identification, la description et la localisation des données (ressources) liées à la santé. Cette structure permettrait non seulement de rechercher et de trouver des données (ressources) spécifiques, mais aussi d'établir des relations entre elles.

### *Considérations relatives à la mise en œuvre*

Les métadonnées sont structurées comme une déclaration en trois parties (sujet - relation - objet) concernant une ressource. La cohérence de la description de la ressource est assurée par un ensemble de modèles désignés par des normes. Le schéma de métadonnées contrôle la structure des déclarations qui peuvent être définies (par exemple, l'ensemble des relations

autorisées), tandis que le modèle d'information contrôle la signification (c'est-à-dire la sémantique) des termes de description selon un vocabulaire standard.

Conformément aux recommandations établies par le W3C<sup>1</sup>, la présente lettre suggère d'adopter un schéma de métadonnées, dit *Dublin Core*, pour décrire les données relatives à la santé. L'ensemble de base du schéma contient 15 éléments (relations) liés à la description des ressources, notamment le créateur, le format, la date, les droits, le sujet, l'identifiant, la langue, etc.

Conceptuellement, un modèle d'information correspond à une ontologie linguistique, représentée sous différentes formes : nomenclature, terminologie, taxonomie, classification et (la plus complexe) ontologie formelle. Dans le contexte d'un dépôt de métadonnées pour un SHIS, le choix du modèle d'information approprié (ou standard de santé) dépend d'un certain nombre de critères à remplir, tels que la couverture du domaine ou la disponibilité dans toutes les langues officielles suisses. Une liste non exhaustive des standards de santé comprend la SNOMED CT (ontologie), LOINC (terminologie), ICD11 (ontologie), ICD10, ICF et ICHI (classifications), WHO-ATC (classification), medDRA et openEHR (terminologies).

### *Opportunités et obstacles*

Pour la présente lettre, une analyse approfondie des principes guidant la conception d'un référentiel de métadonnées pour les ressources de données de santé a été menée. En tenant compte de différentes perspectives (techniques, juridiques et d'utilisation), l'analyse a mis en évidence plusieurs éléments clés liés au sujet. Cependant, un projet potentiel de mise en œuvre d'une infrastructure de gestion des métadonnées des ressources de données liées à la santé soulève un certain nombre de questions pratiques qui devraient être abordées par un échange d'idées, d'opinions et de connaissances entre les parties prenantes concernées (fournisseurs de données, développeurs, décideurs politiques, chercheurs en santé). La capacité du projet à atteindre son objectif principal dépend de plusieurs facteurs importants (opportunités et obstacles).

Les opportunités de mettre en œuvre un dépôt central de métadonnées peuvent inclure :

- l'échange potentiel d'expériences et de savoir-faire avec des projets similaires en Suisse;
- l'intérêt potentiel des décideurs du système de santé (politique ou économique) ;
- un point d'entrée unique pour la recherche/récupération de ressources de données liées à la santé;
- possibilité d'identifier les relations sémantiques entre les ressources de données ;
- une capacité accrue des groupes de recherche à partager/accéder/analyser les données;
- la nécessité d'accroître la transparence des ressources disponibles en matière de santé.

Les obstacles à la mise en œuvre d'un dépôt central de métadonnées peuvent être les suivants:

- une contribution faible ou inexistante des parties prenantes au cours du processus de conception du référentiel de métadonnées;
- un soutien faible ou inexistant de la part des propriétaires de données/fournisseurs de données;
- une absence de consensus sur la norme de soins de santé la plus appropriée pour un référentiel de métadonnées;

---

<sup>1</sup> The World Wide Web Consortium (W3C) est une communauté Internationale qui vise à mettre en place certains standards dans le but d'assurer la croissance du Web sur le long terme.

- aucune structure organisationnelle capable de gérer la mise en œuvre et la maintenance de l'infrastructure matérielle/logicielle du projet;
- l'insuffisance des ressources (humaines et financières) pour passer d'un prototype à un système pleinement opérationnel.